A QM/MM refinement of an experimental DNA structure with metal-mediated base pairs

Kumbhar, S ; Johannsen, S ; Sigel, Roland K O ; Waller, M P ; Müller, J

Abstract: A series of hybrid quantum mechanical/molecular mechanical (QM/MM) calculations was performed on models of a DNA duplex with artificial silver(I)-mediated imidazole base pairs. The optimized structures were compared to the original experimental NMR structure (Nat. Chem. 2 (2010) 229-234). The metal metal distances are significantly shorter (0.5Å) in the QM/MM model than in the original NMR structure. As a result, argentophilic interactions are feasible between the silver(I) ions of neighboring metal-mediated base pairs. Using the computationally determined metal metal distances, a re-refined NMR solution structure of the DNA duplex was obtained. In this new NMR structure, all experimental constraints remain fulfilled. The new NMR structure shows less deviation from the regular B-type conformation than the original one. This investigation shows that the application of QM/MM models to generate additional constraints to be used during NMR structural refinements represents an elegant approach to obtaining high-resolution NMR structures.

DOI: https://doi.org/10.1016/j.jinorgbio.2013.03.009

A QM/MM refinement of an experimental DNA structure with metal-mediated base pairs

Sadhana Kumbhar [a], Silke Johannsen [b], Roland K. O. Sigel [b,*], Mark P. Waller [a,*],Jens Müller [c,*]

[a] *Theoretical Chemistry, Organic Chemistry Institute, Westfälische Wilhelms-Universität Münster, Corrensstr. 40, 48149 Münster (Germany)*

[b] *Institute of Inorganic Chemistry, University of Zürich, Winterthurerstr. 190, 8057 Zürich (Switzerland)*

[c] *Institute of Inorganic and Analytical Chemistry, Westfälische Wilhelms-Universität Münster, Corrensstr. 28/30, 48149 Münster (Germany)*

e-mail: roland.sigel@aci.uzh.ch, m.waller@uni-muenster.de, mueller.j@uni-muenster.de

Dedicated to the memory of Professor Ivano Bertini.

Abstract

A series of hybrid QM/MM calculations was performed on models of a DNA duplex with artificial silver(I)-mediated imidazole base pairs. The optimized structures were compared to the original experimental NMR structure (Nat. Chem. 2 (2010) 229–234). The metal···metal distances are significantly shorter (~0.5Å) in the QM/MM model than in the original NMR structure. As a result, argentophilic interactions are feasible between the silver(I) ions of neighboring metal-mediated base pairs. Using the computationally determined metal···metal distances, a refined NMR solution structure of the DNA duplex was now obtained. In this new NMR structure, all experimental constraints remain fulfilled. The new NMR structure shows less deviation from a regular B-type conformation than the original one. This investigation shows that the application of QM/MM models to generate additional constraints to be used during NMR structural refinements represents an elegant approach to obtaining high-resolution NMR structures.

1. Introduction

Nucleic acids are highly versatile macromolecules[1] that – in addition to their biological relevance – can be applied as self-assembling scaffolds for functional molecules in nanotechnology.[2] Their site-specific functionalization can be achieved by introducing metal ions at pre-defined positions via the formation of metal-mediated base pairs.[3-9] In these base pairs, the hydrogen bonds between complementary nucleobases are formally replaced by coordinative bonds to a metal ion located between the nucleobases. While the naturally occurring pyrimidine nucleosides cytidine and thymidine are known to be capable of forming silver(I)- and mercury(II)-mediated base pairs, respectively, numerous artificial nucleosides have been developed for an application in metal-mediated base pairs. These include for example hydroxypyridone,[10] salen,[11] dipicolylamine,[12] 2,2'-bipyridine,[13] imidazole[14] and triazole.[15] In particular, the imidazole–Ag^+–imidazole base pair has been intensively characterized (Scheme 1a): A comparison with other azole-based nucleosides indicated that it should form highly stable metal-mediated base pairs.[16] This was confirmed by an NMR solution structure of a DNA duplex with three contiguous silver(I)-mediated imidazole base pairs (Scheme 1b).[14] Additional investigations showed that neighboring imidazole–Ag^+–imidazole base pairs are formed in a cooperative manner.[17]

Unfortunately, the most interesting part of the NMR structure from an inorganic point of view, namely that of the metal-mediated base pairs, was poorly defined in terms of Ag···Ag distances:[14] The distance between two neighboring Ag^+ ions was reported to vary between 3.79 Å and 4.51 Å for the ensemble of 20 structures. This can be attributed to the fact that an NMR structure relies on experimentally determined inter-proton distances. Obviously, a nucleic acid derivative in which hydrogen bonds (and hence protons) are formally replaced by coordinative bonds lacks hydrogen atoms in its core. Hence, we applied a quantum mechanical/molecular mechanical (QM/MM) approach to further refine the central section of the NMR structure lacking experimental constraints. Over the past years, advances in QM

methods have increased significantly the accuracy and size of problems that may be successfully investigated. However, biomacromolecules still represent a challenge for pure QM methods. Hence, QM/MM modeling is an important approach in the domain of bio-modeling, as it couples the relative strengths of the QM accuracy to the speed of the MM method. Moreover, the hybrid QM/MM methodology does not require force-field parameters for the silver(I)-containing part of the artificial DNA. As the desired structural accuracy is most crucial in the area of the metal ions, the computationally demanding QM methods can be applied in this region, whereas the "standard" sections of the DNA duplex can be computed using MM.

The behavior and accuracy of specific methods within a given region are well studied and understood, however artifacts at the boundary region between the QM and MM regions may introduce errors. Therefore, the specific combination of QM and MM methods, which gives accurate interaction energies and geometries for particular QM methods, needs special consideration including thorough benchmarking studies of the boundary region. Recently we validated a number of QM methods in combination with the well-established force fields, e.g. CHARMM and AMBER to make sure they are compatible.[18]

2. Material and methods

2.1 QM/MM calculations

A series of models for consecutive metal-mediated imidazole–Ag^+–imidazole base pairs was created starting from the original NMR solution structure (Scheme 1b, PDB ID: 2KE8). QM/MM optimizations were performed using the 2-level ONIOM[19] method in the Gaussian09 software package. The AMBER[20] force field was used; however for the non-natural part of the DNA structure, the atom centered partial charges were assigned using QEq (charge equilibrium) scheme.[21] The B97D,[22] BP86[23-27] and B3LYP[28, 29] functionals were used along with the dispersion correction introduced by Grimme in 2004, which increases the accuracy of DFT for describing the weak inter- and intramolecular dispersion interactions.[30] The 6-31+G (d,p) basis set on light atoms (H, C, N, O) was combined with either the WTBS[31, 32] all electron basis set, or the Stuttgart-Dresden basis set with ECP [33] for Ag. The LANL2DZ (Los Alamos National Laboratory 2 double-ζ)[34-37] basis set, which is a combination of ECP and valence double zeta basis set, was also trialed for Ag.

A series of models with increasing size of the QM region were built starting from the experimental NMR structure of the DNA system (Fig. S2). In model I, one artificial base pair was kept in the QM region, and the remaining atoms were placed into the MM region, and the other two artificial base pairs and the terminal natural base pairs from both ends of the DNA were constrained. In model II, two artificial base pairs were kept in the QM region, and remaining atoms were placed into the MM region, and the single remaining artificial base pair and each terminal natural base pair from both ends of the DNA were constrained. The constraints of terminal base pairs were removed for model III, and three artificial base pairs were placed into QM region, and the remaining DNA was in the MM region. In model IV, the whole DNA duplex was relaxed (no geometric constraints applied), with the central five base pairs in the QM region and remaining DNA in the MM region. Once the model systems for

the best original NMR structure (i.e. the first structure out of the ensemble of 20 NMR structures) were successfully built for model III and the gas-phase optimizations were performed (Fig. S2), the procedure was repeated for the remaining 19 structures in the ensemble of NMR structures. In some instances, the geometry conversion failed. Hence, standard deviations given in this publication consider geometrically converged structures only.

An additional set of gas-phase QM/MM optimizations was performed for model III with B97D/LANL2DZ: AMBER level of theory for the ensemble of NMR structures. The gas-phase optimized structures were solvated with explicit solvent using HyperChem 7.52 software[38] (TIP3P water model,[39] equilibrated at 300 K, 1 atmosphere). These solvated systems (Fig. S3) were re-optimized at the B97D/LANL2DZ: AMBER level of theory. Similarly, a set of solution–phase QM/MM optimizations was performed for model IV.

To study the possible effect of counter ions, a set of QM/MM optimizations was performed for model III including Na^+ counterions in the solvated DNA structure. Na^+ ions were placed at ~3.7 Å distance from each phosphorus atom, along the bisector of two pendant oxygen atoms (Fig. S3). Accordingly, the DNA backbone can be considered neutralized in this version of model III.

2.2 Structure calculation of the refined NMR structure

The refined NMR structure was calculated with Xplor-NIH 2.15.0[40] with the same NMR data restraints and the same 200 input structures as applied for the original NMR structure.[14] To include the $Ag^+\cdots Ag^+$ distances obtained from the QM/MM calculations of model III, three additional $Ag^+\cdots Ag^+$ distance restraints were included during the calculation of the refined NMR structure ($Ag^+\cdots Ag^+$ distances between two adjacent Ag^+ ions were set to 3.35 (–0.1; +0.0) Å and between the two outer Ag^+ ions to 6.60 (–0.1; +0.1) Å). The refined NMR structures fulfill all inter-proton distance constraints and also the additional Ag···Ag distance requirements (Table S2). However, a significant loss of planarity was observed for

the imidazole–Ag^{+}–imidazole base pairs during the first rounds of refinement. As a result, the planarity constraints were increased to obtain planar base pairs (as found also in the core region of the QM/MM structures). The twenty lowest-energy structures out of 200 calculated of the refined NMR structures were visualized and analyzed as described previously[14] using the computer programs MOLMOL[41] and 3DNA.[42]

3. Results and Discussion

3.1 QM/MM Refinement

To test the reliability of the functionals and the basis sets, a small model system was built consisting of a single Ag(I)-mediated imidazole base pair extracted from the best NMR structure (PDB ID: 2KE8). This artificial base pair was placed into the QM region and the remaining atoms (sugar moieties replaced by methyl groups) were placed into the MM region (Fig. S1). Different combinations of the B97D, BP86 and B3LYP functionals with LANL2DZ, 6-31+G (d,p), WTBS and SDD basis sets were considered for the QM/MM(AMBER) calculations for the model system. Table 1 summarizes the results of the performance of these functionals combined with different basis sets. The dispersion included functional B97D along with LANL2DZ basis set was chosen for further studies due to the comparatively lower computational effort compared to other methods, and a rather low RMSD (0.52 Å) compared to the experimental NMR structure.

Dimeric models were built consisting of two neighboring imidazole base pairs at the geometries found in the best original NMR structure (without the DNA backbone, sugar moieties replaced by methyl groups, Fig. 1), i.e. the lowest energy NMR structure out of the ensemble of the 20 NMR structures (PDB ID: 2KE8). To investigate the interaction energies between these base pairs, each imidazole base pair was considered as a monomer A and B respectively, while the dimer model itself as a single complex AB. The interaction energy (*IE*) is defined by equation (I), where E_{AB}, E_A and E_B are total single point energies of the dimer AB, monomers A and B, respectively.

$$IE = E_{AB} - (E_A + E_B) \quad \text{(I)}$$

All QM calculations to obtain the *IE* for the dimer models were performed at tB97D/LANL2DZ level of theory.

As can be seen from Table 2, the *IE* is positive in all gas-phase dimer models. Hence, imidazole base pairs without backbone in the gas-phase are found to be repulsive. For the

adenine:thymine base pairs, the *IE* is –16 kcal/mol at the B97D/LANL2DZ level of theory, which confirms the attractive stacking interaction between the natural system even without backbone. Therefore we conclude that the DNA environment is stabilizing the artificial Ag(I)-mediated imidazole base pairs.

The performance of the functionals (B97D and B3LYP) was evaluated for the best NMR structure over an increasing number of the DNA base pairs in the QM region (for model I, II and III). From Table 3 it is clear that the constraints on the model systems (i.e. freezing the base-pairs from the terminal and artificial base pairs in MM region) lead to a smaller RMSD. Accordingly, the RMSD values for the unconstrained DNA system (model III) are higher, indicating fluctuations in the DNA structure, mostly because the terminal base pairs are free to move. The gas-phase optimized structure still maintains its double helical characteristics.

The gas-phase and solution-phase optimized QM/MM structures for the ensemble of 20 NMR structures were compared to the experimentally refined NMR structures (PDB ID: 2KE8). An overlay of the experimental structure and two of the computed ones is given in Fig. 2. Tables 4 and 5 list the resulting RMSD values and the $Ag^{+}\cdots Ag^{+}$ distances for the various model systems, respectively. The RMSD values indicate that the QM/MM-optimized models already agree rather well with the NMR structure. However, one should keep in mind that the experimental structure was the starting point for these QM/MM calculations. Interestingly, the addition of Na^{+} ions for charge compensation did not result in any significant change in the RMSD values or in the calculated metal$\cdots$metal distances. This is in good agreement with another computational study on DNA with imidazole–Ag^{+}–imidazole base pairs, reporting that the structural stability and the electronic properties of the metal-modified DNA are not impacted by the presence of Na^{+} ions.[43]

The average distance between the Ag^{+} ions in model III in the gas-phase is 3.41 Å compared to 3.42 Å in solution-phase (Table 5). This is significantly shorter than what had been reported in the original NMR structure (Fig. 2). The van der Waals radius for silver is 1.72

Å.[44] Accordingly, the distance between the two Ag^+ ions in the calculated structures is slightly below the sum of two van der Waals radii. Hence, this relatively short $Ag^+\cdots Ag^+$ distance indicates the presence of weak argentophilic interactions.

Comparing the QM/MM optimized structures of model III to those of model IV, the $Ag^+\cdots Ag^+$ distances differ to some extent (Table 5). This could be explained by the observation that in model IV the natural base pairs adjacent to the artificial ones region are somewhat distorted towards the Ag^+ ions. If weak binding interactions between atoms of the natural base pairs and the Ag^+ ions are responsible for the deviation from planarity, then the increased $Ag^+\cdots Ag^+$ distance (and hence decreased $Ag^+\cdots Ag^+$ interaction) might be a direct result of these additional interactions. As model IV is the only model including the adenine:thymine base pairs in the QM region, this effect would not be visible in the other calculations. It is interesting to note that a similar effect had been observed in the molecular structure of another artificial metal-mediated base pair, namely that formed from pyridine-2,6-dicarboxylate, Cu^{2+}, and pyridine.[45]

3.2 Comparison of QM/MM structure and experimental constraints

The QM/MM structures (solvated model III with counterions) were checked against the experimental constraints derived from the original NMR data.[14] In this regard, the inter-proton distances of the geometrically converged QM/MM structures were compared with the respective experimentally determined distance ranges. As a result, most sections of the computed structures did not violate any experimental constraint. However, a couple of NMR distance restraints were consistently not fulfilled by the QM/MM structures (Table S3). These violations are located mainly in the part of the duplex that comprises the natural adenine:thymine base pairs, hence in the MM part. An explanation could be the use of different force fields in the QM/MM calculations (AMBER) and the calculation of the experimental NMR structures (CHARMM). Additionally, the lack of experimental constraints in the QM/MM refinement might be responsible for these violations. Many violations involve

thymine CH_3 groups (Table S3). This can probably be explained by the different treatment of methyl groups in the QM/MM and the NMR methods: In the QM/MM structures, the coordinates are precisely computed for all three hydrogen atoms of a CH_3 group. In contrast, the NMR experiments only provide average information on all three hydrogen atoms of each methyl group, which could then lead to the apparent inconsistencies. Finally, Table 6 lists the six violations consistently observed in the section of the DNA duplex comprising the artificial imidazole nucleosides. Hence, only a minor fraction out of 198 distance restraints are not satisfied in the QM/MM structures. The discrepancies in the case of entries 3 and 6 are probably negligible, as they refer to rather long inter-proton distances anyway, indicating that the original NOE might have been the result of spin diffusion. For the remaining entries, the deviations are rather small and acceptable.

3.3 Refined NMR structure

Finally, new NMR structure calculations were performed (using the methodology described previously[14]). In addition to the experimental inter-proton distance constrains, the refined experimental structure is based on three QM/MM-derived Ag···Ag distance restraints of 3.41 Å (Table 5) (for details see Section 2.2). The ensemble of the 20 lowest energy refined NMR structures (Fig. 3C) does not show any violations of the constraints. Hence, the $Ag^+ \cdots Ag^+$ distances found in the QM/MM calculations do not contradict any of the experimental NMR data. By contrast, they help to describe the metal-mediated base pair region in more detail. The average $Ag^+ \cdots Ag^+$ distance in the ensemble of refined NMR structures amounts to 3.45(2) Å (Table S1). This is about 0.65 Å shorter than in the original NMR structures but within the standard deviation identical to the ones of QM/MM model III. The value also suggests that weak argentophilic interactions are possible between neighboring Ag^+ ions, even though in a somewhat borderline case. As can clearly be discerned from Fig. 3, the minor distortion with respect to the ideal B-DNA conformation[46] as observed for the artificial base pair region of the original NMR structures[14] decreases significantly when employing

the additional $Ag^{+}\cdots Ag^{+}$ distance constraints. The ensemble of refined NMR structures (Fig. 3C) comprises a more squeezed and twisted core region. Further analysis of the global base pair parameters helical twist Ω (the twist angle between two successive base pairs) and helical rise h (the distance between two successive base pairs) confirm a better fit with the B-DNA conformation (Fig. 4). Accordingly, the helical rise of adjacent imidazole–Ag^{+}–imidazole base pairs decreases from 4.1(3) Å to 3.4(3) Å (Fig. 4B). It is interesting to note that the helical rise between an artificial base pair and the neighboring natural base pair is more or less the same in the original and the refined NMR structure. In general, the unwinding of the central part of the duplex as derived from the helical twist Ω is less pronounced in the refined NMR structures compared with the original structures (2KE8). Nonetheless, for the two base pair steps between the imidazole–Ag^{+}–imidazole base pairs an unwinding of the DNA duplex is observed. This can probably be attributed to the fact that the angle between the glycosidic bond angle vectors is not identical for the artificial and the natural base pairs. As already suggested previously, this minor structural deviation might represent an interesting target for the development of a small molecule than recognizes this type of metal-mediated base pairs.[14]

4. Conclusions

The artificial base pairs embedded in a DNA environment (i.e. including the DNA backbone and adjacent natural base pairs) are stable. In the absence of the DNA backbone and when these base pairs are treated as isolated system, they are repulsive in nature according to quantum chemical calculations.

When comparing the gas-phase optimizations and the solution-phase optimizations with the original experimental NMR structure, then the metal-modified DNA system is more compact in the QM/MM structures in the region of the artificial base pairs. This is reflected in a significant decrease of the metal···metal distances compared to the original NMR structure. The role of Na^+ counterions (for charge compensation) was found to be negligible in terms of structural RMSD and $Ag^+ \cdots Ag^+$ distances.

The average helical rise between two artificial base pairs was calculated to change from 4.1(3) Å in the original ensemble of NMR structures to 3.42 Å in solution-phase structure (model III), indicating possible argentophilic interactions of the Ag^+ ions. For the refined NMR structure, an average helical rise of 3.4(3) Å and average $Ag^+ \cdots Ag^+$ distances of 3.45(2) Å were determined.

An important aspect of this investigation is that the inclusion of computationally derived distance information in an NMR structure determination can lead to a significant improvement of structural resolution in an area of the biomolecule for which only limited experimental distance constraints are possible. In the past, the inclusion of distance constraints derived from single crystal X-ray diffraction analyses in an NMR structure determination has become widely accepted.[47] Moreover, QM/MM calculations are well-known to complement experimental data in the elucidation of reaction mechanisms of metalloproteins.[48] The present work nicely shows that the results of QM/MM calculations can complement experimentally derived distance constraints for NMR structure determinations.

5. Acknowledgement

Financial support by the DFG (SFB 858 to JM and MW), the Swiss National Science Foundation (Marie Heim-Vögtlin Fellowship to SJ, and project 200020-143750 to RKOS) as well as the Swiss State Secretariat for Education and Research (to RKOS) is gratefully acknowledged. This work was performed within the framework of COST Action CM1105 (RKOS, JM).

6. References

[1] J. Müller, Metallomics 2 (2010) 318-327.

[2] T.J. Bandy, A. Brewer, J.R. Burns, G. Marth, T. Nguyen, E. Stulz, Chem. Soc. Rev. 40 (2011) 138-148.

[3] P. Scharf, J. Müller, ChemPlusChem in press (2013) doi: 10.1002/cplu.201200256.

[4] D.A. Megger, N. Megger, J. Müller, Met. Ions Life Sci. 10 (2012) 295-317.

[5] Y. Takezawa, M. Shionoya, Acc. Chem. Res. 45 (2012) 2066-2076.

[6] A. Ono, H. Torigoe, Y. Tanaka, I. Okamoto, Chem. Soc. Rev. 40 (2011) 5855-5866.

[7] G.H. Clever, C. Kaul, T. Carell, Angew. Chem. Int. Ed. 46 (2007) 6226-6236.

[8] S. Johannsen, S. Paulus, N. Düpre, J. Müller, R.K.O. Sigel, J. Inorg. Biochem. 102 (2008) 1141-1151.

[9] D.A. Megger, C. Fonseca Guerra, F.M. Bickelhaupt, J. Müller, J. Inorg. Biochem. 105 (2011) 1398-1404.

[10] K. Tanaka, A. Tengeiji, T. Kato, N. Toyama, M. Shionoya, Science 299 (2003) 1212-1213.

[11] C. Kaul, M. Müller, M. Wagner, S. Schneider, T. Carell, Nat. Chem. 3 (2011) 794-800.

[12] K. Seubert, C. Fonseca Guerra, F.M. Bickelhaupt, J. Müller, Chem. Commun. 47 (2011) 11041-11043.

[13] N. Megger, L. Welte, F. Zamora, J. Müller, Dalton Trans. 40 (2011) 1802-1807.

[14] S. Johannsen, N. Megger, D. Böhme, R.K.O. Sigel, J. Müller, Nat. Chem. 2 (2010) 229-234.

[15] D. Böhme, N. Düpre, D.A. Megger, J. Müller, Inorg. Chem. 46 (2007) 10114-10119.

[16] J. Müller, D. Böhme, P. Lax, M. Morell Cerdà, M. Roitzsch, Chem. Eur. J. 11 (2005) 6246-6253.

[17] K. Petrovec, B.J. Ravoo, J. Müller, Chem. Commun. 48 (2012) 11844-11846.

[18] S. Kumbhar, F.D. Fischer, M.P. Waller, J. Chem. Inf. Model. 51 (2012) 93-98.

[19] S. Dapprich, I. Komáromi, K.S. Byun, K. Morokuma, M.J. Frisch, J. Mol. Struct. (Theochem) 461-462 (1999) 1-21.

[20] W.D. Cornell, P. Cieplak, C.I. Bayly, I.R. Gould, K.M. Merz Jr., D.M. Ferguson, D.C. Spellmeyer, T. Fox, J.W. Caldwell, P.A. Kollman, J. Am. Chem. Soc. 117 (1995) 5179-5197.

[21] A.K. Rappé, W.A. Goddard III, J. Phys. Chem. 95 (1991) 3358-3363.

[22] S. Grimme, J. Comput. Chem. 27 (2006) 1787-1799.

[23] P.A.M. Dirac, Proc. Royal Soc. (London) A123 (1929) 714-733.

[24] J.C. Slater, Phys. Rev. 81 (1951) 385-390.

[25] S.H. Vosko, L. Wilk, M. Nusair, Can. J. Chem. 58 (1980) 1200-1211.

[26] A.D. Becke, Phys. Rev. A 38 (1988) 3098-3100.

[27] J.P. Perdew, Phys. Rev. B 33 (1986) 8822-8824.

[28] A.D. Becke, J. Chem. Phys. 98 (1993) 5648-5652.

[29] C. Lee, W. Yang, R.G. Parr, Phys. Rev. B 37 (1988) 785-789.

[30] S. Grimme, J. Comput. Chem. 25 (2004) 1463-1473.

[31] S. Huzinaga, B. Miguel, Chem. Phys. Lett. 175 (1990) 289-291.

[32] S. Huzinaga, M. Klobukowski, Chem. Phys. Lett. 212 (1993) 260-264.

[33] http://www.theochem.uni-stuttgart.de/pseudopotentials (2003)

[34] L.E. Roy, P.J. Hay, R.L. Martin, J. Chem. Theory Comput. 4 (2008) 1029-1031.

[35] P.J. Hay, W.R. Wadt, J. Chem. Phys. 82 (1985) 270-283.

[36] W.R. Wadt, P.J. Hay, J. Chem. Phys. 82 (1985) 284-298.

[37] P.J. Hay, W.R. Wadt, J. Chem. Phys. 82 (1985) 299-310.

[38] HyperChem 7.52 release for Windows`; HyperCube, Inc., Gainesville, Florida, USA, http://www.hyper.com.

[39] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, J. Chem. Phys. 79 (1983) 926-935.

[40] C.D. Schwieters, J.J. Kuszewski, N. Tjandra, G.M. Clore, J. Magn. Reson. 160 (2003) 65-73.

[41] R. Koradi, M. Billeter, K. Wüthrich, J. Mol. Graphics 14 (1996) 51-55.

[42] X.-J. Lu, W.K. Olson, Nucleic Acids Res. 31 (2003) 5108-5121.

[43] P.K. Samanta, A.K. Manna, S.K. Pati, Chem. Asian J. 7 (2012) 2718-2728.

[44] A. Bondi, J. Phys. Chem. 68 (1964) 441-451.

[45] S. Atwell, E. Meggers, G. Spraggon, P.G. Schultz, J. Am. Chem. Soc. 123 (2001) 12364-12367.

[46] R. Chandrasekaran, S. Arnott, J. Biomol. Struct. Dyn. 13 (1996) 1015-1027.

[47] J. Müller, A.A. Lugovskoy, G. Wagner, S.J. Lippard, Biochemistry 41 (2002) 42-51.

[48] L. Rulíšek, U. Ryde, Coord. Chem. Rev. 257 (2013) 445-458.

Tables

Table 1

RMSD (in Å) between best original NMR structure (PDB ID: 2KE8) and QM/MM (AMBER) optimized model system shown in Fig. S1.

Functional	LANL2DZ	6-31+G (d,p) & WTBS	6-31+G (d,p) & SDD
B97D	0.52	1.17	2.21[a]
BP86	1.18	1.16	4.68[a]
B3LYP	1.17	1.17	3.42[a]

[a] Indicates that geometry convergence failed.

Table 2

Interaction energies *IE* of the gas-phase dimer model of two neighboring metal-mediated imidazole base pairs at the B97D/LANL2DZ level of theory.

	IE / kcal mol^{-1}
NMR structure	35.27
Gas-phase optimized	32.33
Solution-phase optimized	32.37
adenine–thymine	–16.71

Table 3

RMSD (in Å, all atoms) of the gas-phase QM/MM optimized structures of the best NMR structure compared to the original best NMR structure (PDB ID: 2KE8). The system indicates an increasing number of DNA base pairs in the QM region.

System	B97D/LANL2DZ: Amber	B3LYP/LANL2DZ : Amber
I	0.96	0.97
II	0.95	0.96
III	2.97	2.25

Table 4

Averaged RMSD values in (Å, all atoms except water) between the QM/MM optimized structures and the respective original NMR structures (PDB: 2KE8).

Gas-phase, model III	Solvated, model III	Solvated with counterions, model III	Solvated, model IV
3.7 ± 1.1	3.4 ± 0.9	3.6 ± 1.3	3.5 ±0.3

Table 5

Average metal···metal distances (in Å) for the various systems. The notations 1–2, 2–3 represent the metal···metal distance of 1st to 2nd and 2nd to 3rd Ag^+ ion, respectively.

Systems	1–2	2–3
NMR	4.10 ± 0.19	4.16 ± 0.19
Model III gas-phase	3.41 ± 0.02	3.41 ± 0.03
Model III solution-phase	3.42 ± 0.04	3.42 ± 0.04
Model III solution-phase + counterions	3.42 ± 0.04	3.41 ± 0.04
Model IV solution-phase	3.49 ± 0.20	3.68 ± 0.14

Table 6

Comparison of inter-proton distances derived from the QM/MM structures (model III) with the respective constrained distance ranges from the experimental NMR structure. Only violations found in the section of the DNA duplex comprising the artificial imidazole nucleosides are listed. For a complete listing, see Table S3.

Entry	Residue 1	Atom 1	Residue 2	Atom 2	computed distance (QM/MM) / Å	lower limit (experimental) / Å	upper limit (experimental) / Å
1	8	H2''	9	H5	2.40(9)	2.8	6.2
2	9	H2'	10	H5	2.5(3)	2.8	6.2
3	9	H5	10	H5''	8.1(2)	3.8	7.2
4	24	H1'	25	H5	4.82(8)	1.6	4.7
5	25	H2''	26	H5	2.5(1)	2.8	6.2
6	26	H5	27	H5''	8.3(2)	3.8	7.2

Figure legends

Scheme 1: a) Representation of an imidazole–Ag^+–imidazole base pair; b) sequence of the DNA duplex under investigation, including nucleotide numbering scheme.

Fig. 1. Gas-phase dimeric model of two neighboring metal-mediated imidazole base pairs; here gas-phase optimized model II (based on the best original NMR structure) is shown.

Fig. 2. Overlay of the best NMR structure (PDB: 2KE8) (in blue), the respective gas-phase optimized structure (in red) and the respective solvated-optimized structure (in green, water excluded for clarity) at B97D/LANL2DZ:AMBER level of theory for model III.

Fig. 3. Comparison of the metal–mediated base pair region: (A) ensemble of the 20 lowest energy original NMR structures[14], (B) ensemble of the 12 geometrically converged structures of the QM/MM-computed model III, and (C) ensemble of the 20 lowest energy refined NMR structures. Depicted are the three imidazole–Ag^+–imidazole base pairs (dark grey) together with the four adjacent natural base pairs (light grey). The structures are superpositioned in the artificial base pair region. The top row clearly indicates the different Ag···Ag distances, the bottom row (rotated by 90° with respect to the upper one) emphasizes the structural differences at the border between QM and MM region. This figure was prepared with MOLMOL.[41]

Fig. 4. Comparison of selected global base pair parameters of the original (grey) and refined (orange) 20 lowest-energy NMR structures. A) Helical twist (i.e. the twist angle between two successive base pairs), B) helical rise (i.e. the distance between two successive base pairs).

Both global parameters are based on C1'-C1' vectors. Absolute values are given on the scale on the right, relative values with respect to an average B-DNA[46] are given on the left.

Scheme 1

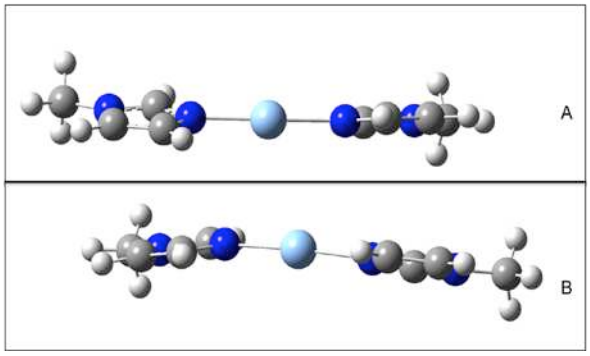

Fig. 1

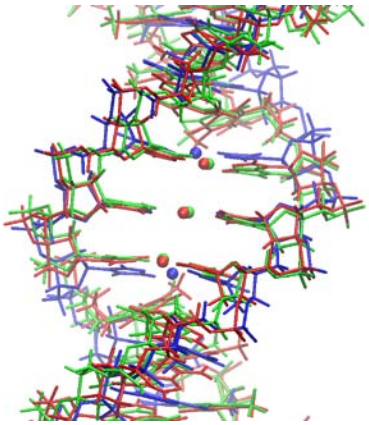

Fig. 2

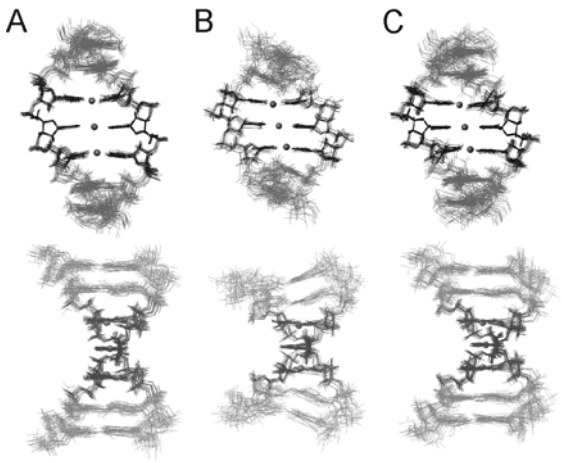

Fig. 3

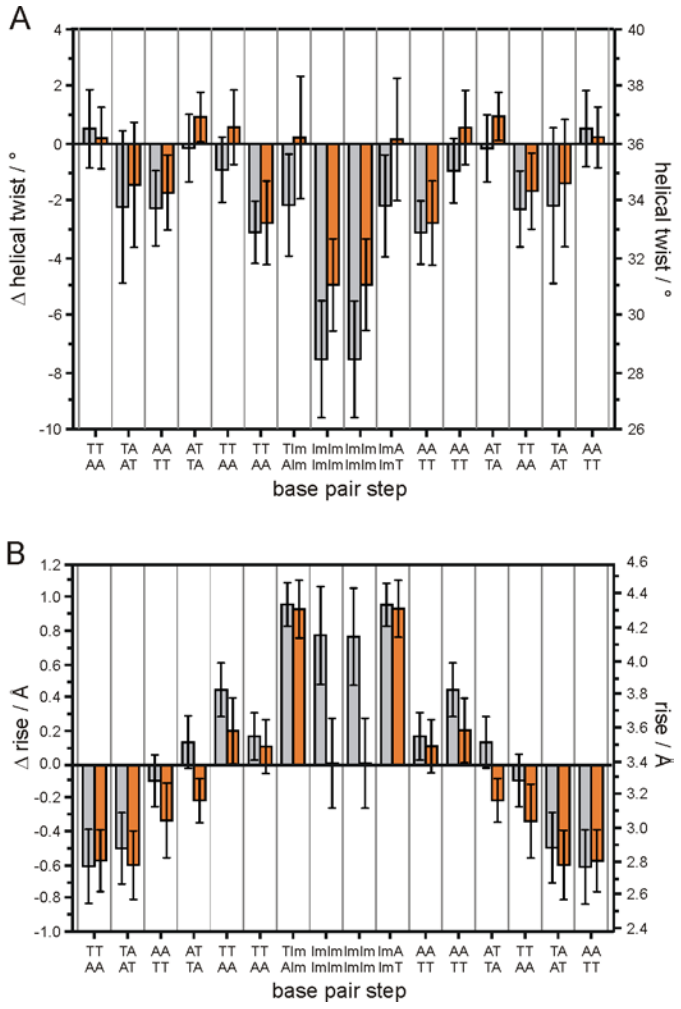
A
Δ helical twist / °
helical twist / °
base pair step
B
Δ rise / Å
rise / Å
base pair step

Fig. 4